



US-CMS Facilities

Ian Fisk
March 16, 2005



US-CMS Facilities



US-CMS Currently Supports and Operates

- Tier-1 Center at FNAL
- 3 Prototype Tier-2 Centers (University of California, San Diego; University of Florida, and Caltech)
 - Ramping this year to seven production facilities
 - Adding University of Wisconsin, MIT, Purdue, and University of Nebraska

The sum of these make up the facility contribution to International CMS Computing

- The US is expected to deliver around 30% of the overall CMS capacity

These facilities also represent the US-CMS facility contributions to the Open Science Grid

- Arrangements for opportunistic use by OSG participants
 - US-ATLAS, LIGO, SDSS, Biology Applications, CS Validations
- Previously we have also made more guaranteed allotments to other VOs for specified periods of time



US-CMS OSG Contributions (1/3)



In addition to the physics computing infrastructure, US-CMS has a number of items we contribute to OSG Development

- ➔ Storage Resource Manager (SRM)
 - FNAL is one of the primary SRM development institutions.
 - US-CMS Currently Supports 0.5FTE and we will grow to 1.5FTE of supported effort this week
 - In addition to performance and robustness improvements, FNAL developers are working on advanced functionality
 - Space Reservation is an active area throughout the spring
 - Improved Authorization should be released soon
 - Working on deployment of SRM-dCache to US-CMS Tier-2 centers
 - Validated installations at UCSD, UFL, Caltech and soon to add UW and Purdue
 - FNAL has been considering providing small scale, high reliability mass storage to smaller OSG Virtual Organizations through SRM
 - Potentially valuable service to communities that require the reliability but cannot afford the operations and expertise associated with automated tape storage.



Cont.

- Virtual Organization Privilege Management
 - Through a joint effort with PPDG common projects and US-ATLAS, Facility developers for US-CMS have been working on improving fine grained authorization
 - Generate extended attribute certificates in VOMS, Parse the certificates by a Globus gatekeeper callout, make a map decision based on the attributes
 - Also permits us to reproducibly map users to a pool of VO Unix accounts
 - Delivered to the OSG Integration Testbed in February
 - Still in beta testing, but successfully deployed at many test sites
- Virtual Organization Management System Extension (VOX)
 - US-CMS in cooperation with iVDGL and FNAL-CD has contributed the VOX project
 - Improve registration and administration of VOMS
 - LCG contribution also



OSG Contributions (3/3)



Number of new projects are about to ramp up

- ➔ FNAL just filled two grid development positions that US-CMS is supporting. New effort arriving this week and more in early May
 - The first should provide development effort to improve auditing and accounting
 - The second is expected to contribute to integration and testing



OSG Operations



In addition to the Facility Infrastructure and Development roles, US-CMS has an operations role within OSG

- ➔ OSG is adopting an operations model with many contributions
 - There is a thin central operations group for tracking and coordination
 - VO Operations
 - US-CMS has a role to support CMS activities and the Tier-1 and Tier-2 centers
 - Service Operations
 - Support for the services contributed to OSG

US-CMS is working to improve our ability to provide support for the distributed centers

- ➔ Providing some central support
- ➔ Improving coordination, configuration, remote support with the Tier-2

Service Challenges and Integration work are improving the supportability of the contributed services.



Service Challenge and Plans



SCI was extremely useful and allowed us to move forward quickly

- ➔ It showed that the currently deployed system has a high degree of usability.
- Using production systems SCI demonstrated a 10x higher throughput (25 TB/day WAN) than prior use in fairly realistic deployment

Many problems exposed due to high number of transfers and high rate. Problems were fixed, or new features added or system redesigned, before proceeding with tests

Rate results:

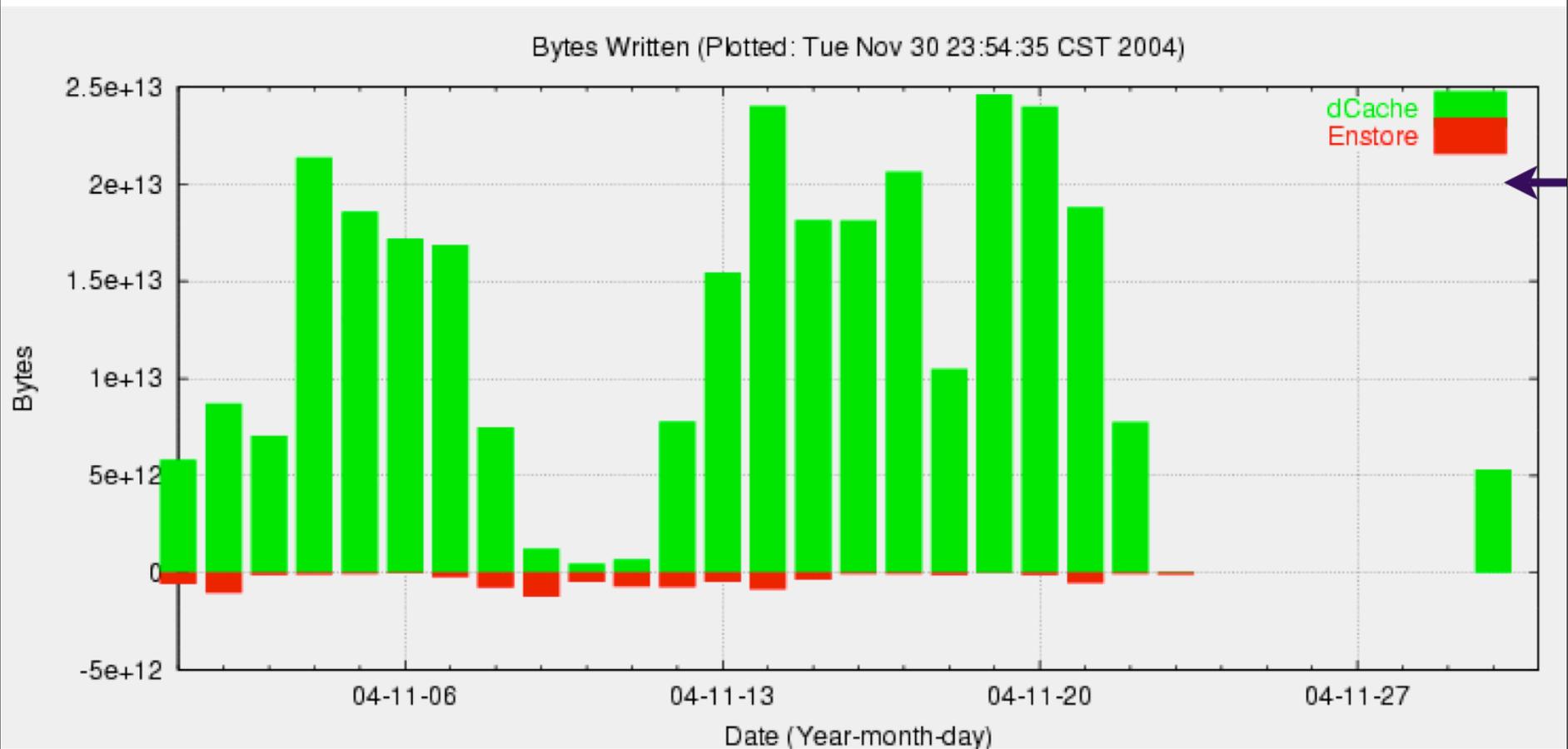
- ➔ Using a fully integrated system, rate - 300 MB/sec, CERN to FNAL, disk to disk
- ➔ Using dcache-java-class gridftp script - 500 MB/sec, no disk at FNAL
- ➔ Using dcache-java-class gridftp script - 400 MB/sec, to dCache disk

Demonstrated our ability to efficiently use the high performance research network

- ➔ Also demonstrated our need for a high performance production link



Data written in SCI by srmcp transfers in November 2004.





Service Challenge II (T0-T1 Interactions)

Continue the goal of making transfers robust and reliable,

- ➔ understand and fix problems rather than just restart

Only use SRM managed transfers to FNAL dCache pools using 3rd party gridftp transfers.

- ➔ No special/contrived transfer scripts.

Continue to use the deployed CMS dCache infrastructure and the Starlight network links. Real data transfers have to coexist with users, service challenge should do so as well. This exposed many bugs in SCI.

Sustain 50 MB/s to tape from CERN to FNAL

Transfer ~10 MB/s of user data and ~40MB/s of fake data

- ➔ Use PhEDEx for both user and fake data transfers - Exercising PhEDEx and making it a robust tool is a goal of SC2

Plan to participate in of 500 MB/s SRM managed transfers to CMS's resilient or volatile pools over the Starlight network links.



Engaging Tier-2 Service Challenge II



We have competent Tier-2 computing centers that want to access CMS data

- ➔ We want to shake out components that we are delivering to OSG

USCMS Tier 1 to Tier 2 Complement of SC2

- ➔ Plan on delivering these data sets to them as part of this challenge also using PhEDEx
- ➔ Initially expect low rate to each, to 3-4 Tier 2 sites
 - End-to-end functionality test of many components
 - Rate can grow as Tier 2 can accept data
- ➔ Sites have installed the PhEDEx daemons.
- ➔ Tier 2 sites do not have MSS, only disk cache, relatively cheap disk. The Tier 1 to Tier 2 operations are being investigated.
 - The disk space is limited, so we will be flushing the Tier-2 disks with some regularity
 - This is reasonable test of CMS computing model elements to treat the Tier2 disk as cache



Service Challenge III



Service Challenge III - stated Goal by LCG:

- ➔ Achieve 50% of the nominal data rate
 - Evaluating tape capacity to achieve challenge while still supporting the experiment
- ➔ Add significant degrees of additional complexity
 - file catalogs
 - OSG currently has only experiment supported catalogs, CMS has a prototype data management system at the same time, so exactly what this means within the experiment needs to be understood
 - VO management software
 - Anxious to reconcile VO management infrastructure in time for the challenge. Within OSG we have started using some of the more advanced functionality of VOMS that we would clearly like to maintain.
 - Uses the CMS's offline frameworks to generate the data and drive the data movement.
- ➔ Requirements are still under negotiation